United Nations ECE/CES/2018/14/Add.13

Economic and Social Council

Distr.: General
15 May 2018

English only

Economic Commission for Europe

Conference of European Statisticians
Sixty-sixth plenary session
Geneva, 18-20 June 2018
Item 11 (a) of the provisional agenda
Programme of work of the Statistics subprogramme of the United Nations Economic Commission for Europe
Reports on the work of the Conference of European Statisticians, its Bureau and Teams of Specialists

Implementation of the United Nations Economic Commission for Europe Statistical Programme 2017

Note by the secretariat

Addendum

Report of the Workshop on Statistical Data Collection 'Overcoming Challenges through Modernisation'

Summary

The document presents the key outcomes of the Workshop on Statistical Data Collection "Overcoming Challenges through Modernisation" which took place from 10 to 12 October 2017 in Ottawa, Canada. The meeting was organized following a decision of the Conference of European Statisticians in 2017.

The report is submitted to the Conference of European Statisticians for information.

GE.18-07762(E)
1807762

Please recycle

I. Attendance

1. The Workshop on Statistical Data Collection: Overcoming Challenges through Modernisation was held in Ottawa, Canada, from 10-12 October 2017. It was attended by 70 participants (64 representatives and 6 observers). They represented the statistical offices of Australia, Canada, Curaçao, Denmark, Estonia, Finland, France, Germany, Ireland, Israel, Italy, Mexico, Montenegro, Netherlands, Norway, Philippines, Poland, Portugal, Russian Federation, Slovenia, Switzerland, United Kingdom of Great Britain and Northern Ireland and United States of America as well as the International Labour Organization (ILO), the International Monetary Fund, the United Nations High commissioner for Refugees and World Bank. The Goldsmith College of the University of London, Washington State University, and the NPD Group were also represented

2. The Workshop was hosted by Statistics Canada and was organized under the responsibility of the High-Level Group Modernisation Committee on Statistical Products and Sources. The Organizing Committee consisted of Barteld Braaksma (Statistics Netherlands), Frances Comerford (Central Statistical Office Ireland), Bilal Kurban (Turkish Statistical Institute), Lise Rivais (Statistics Canada), Evelyn Ruppert (Goldsmith College of University of London), Paulo Saraiva dos Santos (Statistics Portugal) and Niki Sylianidou (Eurostat).

II. Organization of the meeting

3 The workshop was chaired by Mrs. Lise Rivais from Statistics Canada. Mr. Stéphane Dufour, Assistant Chief Statistician of Statistics Canada, opened the workshop and welcomed the participants. Mr. Don A. Dillman from Washington State University provided a keynote address.

4. The agenda included the following substantive topics:

(a) Alternative sources and ways of data collection to reduce survey burden;

(b) Modernisation of data collection;

(c) Use of new technologies, skills and mind-set for data collection;

(d) Management of data collection;

(e) Communication with Respondents.

5. In total, two keynote addresses and thirty presentations were given within the five sessions. Each session was concluded with a small group discussions focusing on lessons learned and on activities in data collection to eliminate, reduce, increase and create. The outcomes of the discussions are documented in the annex. A voting round took place to identify the most relevant topics in each of the four action types for future work.

6. All abstracts, papers and presentations from the Workshop on Statistical Data Collection - Workshop are available at the workshops webpages https://statswiki.unece.org/x/cQKoBw.

III. Main conclusions reached at the meeting

7. Key items identified for future work included (by type of action and in order of preference):

A. Eliminate

- Eliminate the use of non-reusable devices;
- Eliminate duplication of efforts by more collaboration;
- Eliminate issues with access to alternative data by removing legal and other barriers;
- Eliminate stove pipe and silo structures and create standardised processes.

B. Reduce

- Reduce the length and complexity of questionnaires without losing information;
- Reduce the time to make a decision and accept better the risks of innovation such as new collection strategies;
- Reduce the use of surveys by using administrative data and other sources;
- Reduce the number of data silos towards a more integrated data collection framework.

C. Increase

- Increase investment in human resources (capacity building) to have the capacity/capability for technological change;
- Increase international cooperation in innovative projects;
- Increase the sharing of code and stimulate open source;
- Increase providing personalized feedback to respondents.

D. Create

- Create modular solutions to reduce the number of surveys;
- Create interactivity and dialogue with the respondents (e.g. 'gamification' of surveys);
- Create dedicated teams to research the use of alternative sources such as big data;
- Create communication based on behavioural and psychological testing;
- Create systems that can deal with multi-mode data collection.

Annex

Summary of discussions on substantive topics

A. Topic (i): Alternative sources and ways of data collection to reduce survey burden

1. This session was organized by Lisa Rivais (Statistics Canada) and Niki Stylianidou (Eurostat). In included the following presentations:

- United States Census Bureau: Evaluating the use of publicly available building permit data to compliment data collection and analysis of construction surveys;
- Istat (Italy): Integration of administrative and survey data to reduce respondent burden: the Italian experience in the field of agriculture statistics;
- World Bank: Surveying informal enterprises: applying stratified adaptive cluster sampling using CAPI with implementation and monitoring tools;
- United States Census Bureau: Reducing survey burden through third-party data sources;
- Statistics Netherlands: Dissolving Questionnaire Borders with Technology;
- Statistics Portugal: Using Administrative Data for Statistical purposes - Statistics Portugal Experience;
- Statistics Netherlands: Centre for Big Data Statistics.

2. The following points were identified by the participants as lessons learned:

- It takes time to process the structured data coming to statistical organization - processing time and challenges should not be underestimated;
- Administrative data can be used to validate statistics but also the reverse;
- Use more administrative data;
- Apps can be useful.

3. The following actions were identified for future work:

(a) Eliminate:

- some of the policy and legal barriers currently in place restrain access to the information;
- rationalize/reduce/eliminate redundancies;
- increase use of partial substitution of admin data rather than full substitution aiming to eventually get full substitution;
- redundancy, especially for business surveys;
- duplication;
- statistics which may be inconsistent, obsolete, in duplication;
- complexity of data requests;
- paper.

(b) Reduce:

- legal requirements in terms of confidentiality of the information;
- consolidating data requests (could also eliminate);

- duplication of collecting the "same" data from multiple sources ... merging surveys together;
- number of surveys;
- dependency on other organizations;
- cost to programmes;
- burden;
- numbers of primary surveys;
- cost;
- respondents burden;
- time to market.

(c) Increase:

- a culture shift is required in terms of how the collected data is to be used, the quality of the information collected and the process to use this data in official statistics;
- the possibility that the current historical series may not be maintained;
- comparability issues and methodological validation required in using new data - robustness and quality of the data;
- the possibility that the concepts may differ with the use of structured big data
- use of admin data;
- self-response that does not require follow up first;
- personalized feedback for data collection;
- administrative comparability between countries: advantage or disadvantage;
- administrative data and meeting new requirements;
- integration of multiple sources into a single statistics;
- access to big data;
- coherence analysis;
- external partnership;
- collaboration, modernisation;
- trust through outreach, communication, education;
- linking of data, increase richness.

(d) Create:

- new system able to process the structured information coming for administrative sources in a timely manner;
- legislation/agreement for supply and use of alternative sources replacing survey data;
- new data by data linkage, maybe easier than big data initiatives;
- specific teams to research alternative sources such as big data;
- Investment in collaboration with other organizations;
- better experience for respondents;
- methodology to relevance of statistics;
- branding of statistics – awareness;
- harmonization (variables, questionnaires);
- added value for respondents;
- Innovation in data analysis, visualization.

B. Topic (ii): Modernisation of data collection

4. This session was organized by Mr. Paulo Saraiva (Statistics Portugal) and Niki Stylianidou (Eurostat). It included the following presentations:

- ONS (United Kingdom of Great Britain and Northern Ireland): Transformation of UK Business Data Collection;
- CSO (Ireland): Modernisation of Business Statistics data collection;
- Statistics Netherlands: Phoenix: Redesign of the data collection process and systems;
- ONS (United Kingdom of Great Britain and Northern Ireland): Transforming social survey data collection in the UK – opportunities and challenges;
- Istat (Italy): A standard procedure for the optimization of survey questionnaires design;
- Statistics Denmark: Standardized and modernized data editing in Statistics Denmark;
- International Monetary Fund: Moving to Automated Collection using SDMX.

5. The following actions were identified for future work:

(a) Eliminate:

- some administrative constraint in the decision making process;
- number of silo;
- redundant system;
- unique definitions for similar concepts;
- legacy system;
- duplication of effort by more collaboration;
- stove pipe silos and create standardise processes;
- fear of using non-traditional modes such as social medias.

(b) Reduce:

- number of teams involve in decision process;
- number of owner of various system;
- the number of silo data to a more integrated framework;
- the number of different questionnaire designs (variation between questionnaires);
- over-editing in questionnaire;
- business contacts by integrating surveys;
- response burden;
- length and complexity of questionnaires without losing information;
- time to make a decision and accept better the risks of innovation such as new collection strategies;
- Reduce traditional collection modes with more non-traditional modes i.e. social medias.

(c) Increase:

- realistic funding assessment;
- the importance of higher organisational goal over survey specific needs;
- complexity of modernisation to higher management;
- resistance to change;
- alignment of the organizational structure to match new systems/processes;
- better and specific documentation for legacy systems. Difficult to transition from old systems to new without understanding what old systems did;
- usage of modular approach;
- sharing code, open source;

- finding new ways to engage respondents for different survey topics and various segments of the population;
- respondent relations in social and business surveys are different - e.g. continuity and trust with business surveys;
- how and when to prefill questionnaires - raw or edited data? Do we need guidelines?
- creation of structure for peer support across different NSI with different capacities;
- digital communication through all stages of processes;
- going beyond conventional cognitive testing with usability testing.

(d) Create:

- new governance to align with new system and process;
- awareness of the complexity of modernisation process;
- a system that can deal with multi-mode collection;
- pre-defined quality standards to prevent over editing;
- better standards for documentation;
- allocation of budget and resources for innovation;
- the culture;
- develop and implement generalized processing: data collection and common approaches and guidelines applicable to all surveys;
- interactivity and dialogue with the respondents i.e. "gamification" of surveys.

C. Topic (iii): Use of new technologies, skills and mind-set for data collection

6. This session was organized by Bilal Kurban (Turkish Statistical Institute) and Barteld Braaksma (Statistics Netherlands). It included the following presentations:

- Estates (Germany): The impact of ongoing technical developments on the entire data production – from concepts, programming to testing and data collection;
- MONSTAT (Montenegro): Agricultural Census in Montenegro – methods of data collection;
- World Bank: Survey Solutions - an integrated system of data collection, survey management and data aggregation;
- Statistics Netherlands: Questionnaire design for smartphones (and tablets) – an on-going journey from mobile unfriendly to mobile friendly to mobile optimization;
- Goldsmiths, University of London: A citizen data app: reflections on a collaborative practice;
- World Bank: Surveying Informal Enterprises: Applying Stratified Adaptive Cluster Sampling using CAPI with Implementation and Monitoring Tools.

7. The following actions were identified for future work:

(a) Eliminate:

- stigmatisation of failure in research;
- overlap between innovation and production;
- working towards a gold standard and accept a workable solution;
- not being stuck in our old ways;
- roadblocks to creative thinking without being too critical;
- long questionnaires which are not suitable for mobile collection.

(b) Reduce:

- time frame for decision making (such as funding) in innovation projects;
- risk aversion;
- single mode surveys;
- costs;
- use of surveys using administrative data or other sources;
- amount of survey data in our statistics - new mix of modes and data sources.

(c) Increase:

- expertise in new technologies;
- international cooperation in innovative projects;
- investment in human resources (capacity building) to have capacity/capability for technological change;
- pace among others partners of the current collection challenges;
- usability to have a better experience for respondent;
- alternative modes to align with respondent requirements;
- communication using multi-mode package to increase response rates. Address differences in communications via apps which calls for person to download which is different from a link from a web platform;
- new method requirement to work with technology constraints;
- new methods for new technologies e.g. design for mobile devices;
- different concepts, definitions, categories of data taking into account the mode e.g. big data;
- convincing subject matter to use new technology to optimize collection.

(d) Create:

- an agile governance in and for innovations;
- structure solely dedicated to transformation;
- modular solutions to reduce numbers of surveys;
- administrative centralized and standardized databases;
- new infrastructure to address new technology;
- profiling functionality in questionnaires;
- apps to measure things that are difficult to measure;
- environment conducive to innovation/experiment;
- community of best practices for modern survey design (mobile /online questionnaire);
- design for Smart Phone should tested;
- new relationship and collaboration with society and individuals;
- redesign focus from survey to mode or interface;
- prepare for susceptibility to risks with different digital technologies such as breakdowns. Prepare communications and backup plan;
- new methods for new technologies.

D. Topic (iv): Management of data collection

8. This session was organized by Paulo Saraiva (Statistics Portugal) and Evelyn Ruppert (Goldsmith, University of London). It included the following presentations:

- Statistics Canada: Active Management Framework to Monitor and Manage Data Collection;

- Statistics Estonia: Modernizing data collection preparation, management, data correction. Integration of data collection infrastructures and technologies to optimize processes and resources;
- Statistical Office of the Republic of Slovenia: Intersectoral partnership and outsourcing of data collection in the Statistical Office of the Republic of Slovenia;
- Statistics Canada: Statistics Canada's Experiences to Manage Responsive Collection Design for Household Surveys;
- INEGI (Mexico): Planning the 2020 Census in Mexico: perspective and challenges for data collection;
- Statistics Poland: CORStat as a tool for managing surveys and interviewers.

9. The following actions were identified for future work:

(a) Eliminate:

- Operational constraint with modern tools
- Manual processes
- Monolithical approach.
- Non formal management - be more proactive
- Bad behaviours with different monitoring techniques such as GPS
- Use of non-reusable devices / platforms
- Over collection once target has been met

(b) Reduce:

- Complexity of monitoring tool
- Operational burden by sharing workload (between surveys for example)
- Paper management burden
- Use of incentives (gifts) which are causing some sample groups to be over represented. (i.e. lower incomes over represented).
- Technology barriers by leaving legacy systems
- Technical difficulties; limitations of systems (the ones we put on ourselves).
- Reduce complexity of changing mode during collection
- Reducing the risks on non-bias by introducing efficiencies
- Focus on hard to get respondents unless they are high priority
- Over management and micro management of field force.

(c) Increase:

- Importance of using GIS data
- Field logistics complexity in large scale operation
- Use of integrated case management to transfer/communicate between collection modes
- Use of population registers
- Modularity of case management systems.
- A framework to monitor collection
- Best practices and challenges
- International collaboration in regard to different collection systems, collection management, training, etc.
- Feedback to field interviewers on progress
- Repurposing digital devices with other data collection or school.

(d) Create:

- International cooperation in developing complex systems with similar requirements

- Efficient communication strategy between all partners in NSI
- Tools for efficient resource allocation
- Partnership to share resources (including physical resources)
- Automation of allocation of work
- Machine learning
- Create synergy between data collection of social & business statistics
- Create automatics allocation of work in the field
- Create a more flexible communication system in case management
- Specific data collection coverage target by priority
- Management system on business statistics.

E. Topic (v): Communication with respondents

10. This session was organized by Frances Comerford (CSO, Ireland). It included the following presentations:

- Statistics Denmark: Optimization of data collection from enterprises by use of nudging initiatives;
- Central Statistical Office Poland: Communication with respondents in business surveys conducted by Reporting Portal;
- Statistics Finland: How to motivate and catch hard-to-reach groups in household surveys;
- Statistics Estonia: Activities to reduce response burden: pre-filling of questionnaires, development of questionnaires and electronic data submission environment eSTAT based on feedback from respondents and data processing phase;
- ONS (United Kingdom of Great Britain and Northern Ireland): Using Behavioural Science to influence respondents … and ourselves.

11. The following actions were identified for future work:

(a) Eliminate:

- traditional approach
- barriers to communication with respondents by ensuring multiple contact options (i.e. phone, email, etc.)
- complex survey forms
- one dimensional approach of population
- extensive - complicated communication with respondents.

(b) Reduce:

- use of non-monetary incentives
- time lag for collection periods
- unnecessary information (keep it simple) i.e. too much official information that respondents do not read
- simplify communication such as a letter to respondent
- raise good respondent experience
- personal support by providing more interactive digital support.

(c) Increase:

- awareness of collection activities between intercensal period
- user experience interest on the collection portal

- use of behavioural economics
- outreach to schools or difficult to reach respondents
- opportunities for respondent feedback
- self-administrated customer service.
- personalize the collection of data on large enterprises
- knowledge about key contact information
- good respondent experiences.

(d) Create:

- reports to return to respondents with relevant data
- use of social media
- portal for business respondents
- personalized feedback data/products to respondents to incentivize
- more effective respondent communication based on behavioural theory and respondent behaviour
- opportunity to test different ways of communication
- dedicated team for the enterprise surveys
- user experience measures- index
- new non-monetary incentives
- instant and personalized feedback.

F. Lessons learned

12. Across the sessions, the following points were identified as lessons learned:

- It takes time to process the structured data coming to a statistical organization - processing time and challenges should not to be underestimated;
- Administrative data can be used to validate statistics but also the reverse;
- Use more administrative data;
- Apps can be useful;
- Investigate use of and access to Big Data;
- Administrative data can be used to validate statistics, but also the other way around;
- There is a discrepancy between legislation and public perception;
- Parallel development and production is not ideal;
- Adopt one step at a time;
- Systems cannot be fully standardized;
- Feedback to respondents and the measures of results of the benefits of these feedback, is a good way to improve relationships
- Great difference between countries and diversity on how they will achieve the same goals;
- Design for mobile collection first, then move on to design for other devices;
- Question the use of non-monetary incentives.